



Learning efficient visual search for stimuli containing diagnostic spatial configurations and color-shape conjunctions

Eric A. Reavis^{1,2,3} · Sebastian M. Frank^{1,4} · Peter U. Tse¹

© The Psychonomic Society, Inc. 2018

Abstract

Visual search is often slow and difficult for complex stimuli such as feature conjunctions. Search efficiency, however, can improve with training. Search for stimuli that can be identified by the spatial configuration of two elements (e.g., the relative position of two colored shapes) improves dramatically within a few hundred trials of practice. Several recent imaging studies have identified neural correlates of this learning, but it remains unclear what stimulus properties participants learn to use to search efficiently. Influential models, such as reverse hierarchy theory, propose two major possibilities: learning to use information contained in low-level image statistics (e.g., single features at particular retinotopic locations) or in high-level characteristics (e.g., feature conjunctions) of the task-relevant stimuli. In a series of experiments, we tested these two hypotheses, which make different predictions about the effect of various stimulus manipulations after training. We find relatively small effects of manipulating low-level properties of the stimuli (e.g., changing their retinotopic location) and some conjunctive properties (e.g., color-position), whereas the effects of manipulating other conjunctive properties (e.g., color-shape) are larger. Overall, the findings suggest conjunction learning involving such stimuli might be an emergent phenomenon that reflects multiple different learning processes, each of which capitalizes on different types of information contained in the stimuli. We also show that both targets and distractors are learned, and that reversing learned target and distractor identities impairs performance. This suggests that participants do not merely learn to discriminate target and distractor stimuli, they also learn stimulus identity mappings that contribute to performance improvements.

Keywords Perceptual learning · Visual search

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13414-018-1516-9>) contains supplementary material, which is available to authorized users.

✉ Eric A. Reavis
ereavis@ucla.edu

Sebastian M. Frank
sebastian_frank@brown.edu

Peter U. Tse
peter.u.tse@dartmouth.edu

¹ Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA

² Present address: Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA 90024, USA

³ Present address: Desert Pacific Mental Illness Research, Education, and Clinical Center, Greater Los Angeles Veterans Affairs Healthcare System, Los Angeles, CA 90073, USA

⁴ Present address: Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, RI 02912, USA

Perceptual learning

To survive and prosper in the world, humans and other animals must process and act upon meaningful sensory input as quickly and accurately as possible. Stimuli that are meaningful in one environment, however, may be meaningless or non-existent in many others. Thus, the ability to optimize processing of sensory information to the particular environment in which an organism finds itself is important.

People and animals do indeed have a capacity to optimize perception to fit the environment, and this capacity is called perceptual learning (Doshier & Lu, 2017; Gibson, 1963; Goldstone, 1998; Sasaki, Nánéz, & Watanabe, 2010; Watanabe & Sasaki, 2015). Decades of research have characterized perceptual learning in every major sensory modality, and it has been especially well studied in the domain of visual perception (Gold & Watanabe, 2010). Several recent reviews provide comprehensive overviews of the visual perceptual learning literature, which has become quite extensive

(Doshier & Lu, 2017; Maniglia & Seitz, 2018; Watanabe & Sasaki, 2015).

Most perceptual learning studies to date have shown, using various methods, that with repeated exposure, observers become better at detecting or discriminating single task-relevant visual features. For example, seminal studies showed that training improved subjects' processing of the orientation of lines presented in isolation (Shiu & Pashler, 1992) or embedded in a texture-discrimination task (Karni & Sagi, 1991, 1993), the offset of Vernier stimuli (Fahle, Edelman, & Poggio, 1995; Poggio, Fahle, & Edelman, 1992), and the direction of motion of fields of dots (Ball & Sekuler, 1982, 1987; Watanabe, Náñez, & Sasaki, 2001).

Early studies found that such perceptual learning effects were highly specific to the particular features experienced during training and the retinotopic locations where those features appeared. For example, observers improved their perception of particular line orientations at a particular location in retinotopic space, but returned to baseline performance when lines were presented at a different orientation or in a different part of the visual field (Karni & Sagi, 1991; Shiu & Pashler, 1992). Similarly, repeated practice discriminating Vernier stimuli at particular retinotopic locations led to enhanced acuity in offset perception for those stimuli in those locations, but it did not generalize to other retinotopic locations, slightly rotated stimuli, or to a slightly different type of offset stimulus (Fahle, 1997, 2004; Fahle & Morgan, 1996; Poggio et al., 1992). Improvements in motion perception were also found to be specific to the particular retinotopic location and motion direction shown during training (Ball & Sekuler, 1982, 1987; Watanabe et al., 2001). More recent studies, however, have found that these types of single-feature perceptual learning effects can sometimes generalize to new retinotopic locations and feature values outside the trained range when certain task parameters are adjusted (e.g., task difficulty, number of trials, task complexity; Ahissar & Hochstein, 1997; Harris, Gliksberg, & Sagi, 2012; Wang et al., 2016; Wang, Zhang, Klein, & Levi, 2014; Xiao et al., 2008).

Learning complex combinations of visual features

In the real world, meaningful stimuli are often defined by a complex conjunction of features that can occur anywhere in the visual field: stop signs are red and octagonal; grapes are green (or red) and round and densely clustered. Typically, perceptual processing of visual feature conjunctions is slow and inefficient (Treisman & Gelade, 1980). Visual feature conjunction processing has been studied extensively with visual search tasks, in which stimulus processing efficiency can be measured precisely by varying the number of search items and computing the required processing time per item.

One of the most difficult (i.e., typically inefficient) types of feature conjunction search is for stimuli defined by the configural relation between two elements, such as a bisected pattern containing two colors in one arrangement to define a target (e.g., red/green) and in the opposite arrangement to define distractors (e.g., green/red; Sripati & Olson, 2010). Figure 1 contains an example of such a search array.

Longstanding evidence shows that visual search for conjunctive stimuli defined by the spatial configuration of two elements can become more efficient with training. One of the first studies to demonstrate such conjunction learning (CL) trained participants to search a set of bisected squares that were red on one side and blue on the other (Heathcote & Mewhort, 1993). Over the course of two 1-hour training sessions on different days, search efficiencies improved dramatically. Similar results were obtained when target and distractor identities alternated throughout training (i.e., a variable mapping, where participants searched for an oddball which could be either red/blue or blue/red; Heathcote & Mewhort, 1993).

Around the same time, Wang, Cavanagh, and Green (1994) investigated visual search for well-practiced configural conjunction stimuli. Their subjects searched letters formed by different configurations of the same line elements (such as Ns and Zs, or stylized 2s and 5s). Stimuli presented at canonical orientations were processed efficiently, but stimuli presented at alternative orientations (i.e., sideways) were processed inefficiently. Distractor orientation was more critical for efficient search than target orientation.

Another group investigated learning in visual search for stimuli defined by figural cues such as closure and convergence (Sireteanu & Rettenbach, 1995, 2000). Over two training sessions on different days, processing efficiency improved dramatically. In addition, learning transferred almost completely from one type of figural cue to another, and from one retinotopic location to another. Furthermore, learned improvements persisted after a 4½ month delay in training.

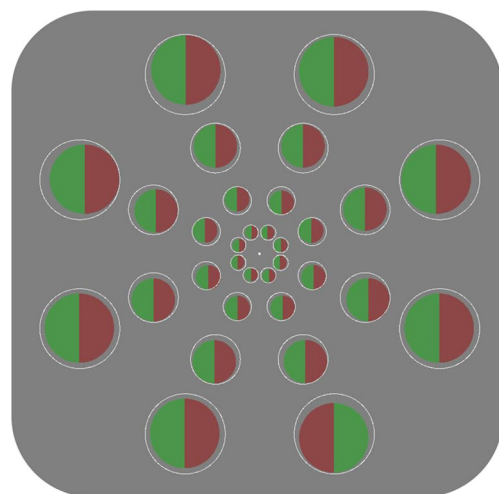


Fig. 1 Example search array (target in ring 4)

Ellison and Walsh (1998) trained participants in a variety of visual feature conjunctions (color/size, color/orientation, oriented configurations) for eight sessions of 900 trials each, on different days. They found that learning occurred for all three types of stimuli. A follow-up experiment showed that transcranial magnetic stimulation to the right parietal cortex disrupted visual search for unlearned feature conjunctions, but not searches for learned conjunctions (Walsh, Ashbridge, & Cowey, 1998).

Meanwhile, Carrasco, Ponte, Rechea, and Sampedro (1998) investigated the effect of distractor homogeneity on learning, using color/orientation and color/position conjunctions. They found that learning progressed more quickly when all distractors shared a feature not present in the target (e.g., when all distractors shared a blue patch but the target contained no blue). However, they found that processing efficiency improved with learning for all of the conditions they tested.

In a series of recent studies, we used fMRI to investigate learning of bisected configural conjunction stimuli in visual search. CL was associated with a distinctive neural signature: As search efficiency improved, the amount of activity in the retinotopic location of the target increased relative to the amount of activity in distractor locations throughout various regions of visual cortex (Frank, Reavis, Tse, & Greenlee, 2014). In other words, maps of retinotopic space throughout the visual cortex contained an activity peak for targets after learning. Surprisingly, this signature of learned target “pop-out” in the retinotopic cortex was still evident after training, even when participants performed an attention-demanding fixation task and the irrelevant learned stimuli were presented in the periphery (Reavis, Frank, Greenlee, & Tse, 2016). The effect was specific to the learned stimuli and did not result from a general improvement in visual-search performance; when participants were tested under the same conditions with a totally different type of conjunction stimuli after training, the neural signature disappeared (Reavis et al., 2016). We also found similar target versus distractor enhancement effects for different types of conjunction stimuli (e.g., motion-defined conjunction stimuli; see Frank, Reavis, Greenlee, & Tse, 2016), and that such effects can persist for years after training (Frank, Greenlee, & Tse, 2018).

Determining what is learned

The existing literature shows that CL can occur relatively quickly for various different types of stimuli and that the learning can be both durable and generalizable under certain circumstances. Indirect evidence suggests that CL might involve learning to ignore distractors, at least in part (Carrasco et al., 1998; Wang et al., 1994). At the neural level, the evidence suggests that CL leads visual-search targets to “pop out” on

retinotopic maps in various visual areas with minimal top-down attention.

Yet it remains unclear what exactly subjects are learning. It is possible that subjects learn to make use of basic image-level features of the search array that indicate target presence or absence (e.g., learning that the appearance of red at particular retinotopic coordinates indicates the presence of a target). It is also possible that subjects are learning to use integrative, high-level properties of the stimuli (e.g., the relational structure of the stimulus configuration or color–shape associations). For example, training might gradually change the tuning properties of units high in the visual hierarchy so that particular feature conjunctions associated with search targets and distractors activate some neurons more selectively (e.g., developing target and distractor “detectors” that respond to one stimulus, but not the other).

These two possibilities are both countenanced by an influential model for understanding perceptual learning and its relationship to visual search: reverse hierarchy theory (RHT; Ahissar & Hochstein, 2004; Ahissar, Nahum, Nelken, & Hochstein, 2009; Hochstein & Ahissar, 2002). According to RHT, ordinary conscious perception primarily reflects the activity of neurons in advanced areas of the cortical visual processing hierarchy. Because the activity of cells in such areas effectively summarizes the contents of a scene by responding selectively to particular objects, faces, and the like, the gist of the scene can be apprehended quickly and easily “at a glance” by assessing the activity of these high-level neurons. By contrast, information encoded at earlier levels of the visual hierarchy about the featural details of the scene (its contours, colors, etc.) is less readily accessible and must be accessed via slower, more effortful mechanisms involving focused attention by a process termed “vision with scrutiny.”

RHT proposes that one major way perceptual learning occurs is that observers gradually attain more direct access to particular units at lower levels of the visual hierarchy where task-relevant information is encoded, making “vision with scrutiny” easier, faster, and more effective over time. Later, RHT posits, continued training gradually changes the tuning of neurons in high-level areas so that they respond more selectively to task-relevant stimuli. In this way, according to RHT, expert observers are able to detect and respond to learned stimuli quickly “at a glance” because their training has effectively created high-level detector units for those stimuli.

In this study, we present a series of experiments that test these two differing (but not mutually exclusive) predictions of RHT as an explanation of CL. In each experiment, participants trained on a CL task for a week. Then, one property of the stimuli was manipulated and performance was measured again. Post-manipulation search performance was compared to performance at the beginning of training to determine if the learning generalized to the new stimuli (i.e., performance was

significantly better than at the beginning of training). Post-manipulation performance was also compared to performance at the end of training to determine if processing of the stimuli was significantly disrupted by the manipulation.

We first describe a preliminary experiment that demonstrates the CL effect and the magnitude of performance disruption that can occur when the learned stimuli are manipulated. Experiments 1a and 1b test the hypothesis that trained participants achieve better “vision with scrutiny” by learning to make use of low-level feature information to perform efficient visual search. Experiments 2 and 3 test the hypothesis that trained participants learn to perceive high-level properties of the stimuli (i.e., particular conjunctions of visual features) efficiently “at a glance.”

For these experiments, we predicted that CL would involve learning to use high-level characteristics of the stimuli and not low-level feature information. This hypothesis was predicated on prior studies of learning in visual search for different types of feature conjunctions, which found learning transfer when low-level features of the stimuli were manipulated (Sireteanu & Rettenbach, 1995, 2000). In addition, a recent study of learning in visual search for color-orientation conjunctions using different methodology found evidence for learning of feature conjunctions rather than learning of individual stimulus features (Yashar & Carrasco, 2016). Furthermore, studies of visual search for real-world stimuli have found that distractors that share high-level features with the target (e.g., object category) can impede search efficiency even when few low-level features are shared between the target and distractors (e.g., Wyble, Folk, & Potter, 2013). Similarly, distractors from object categories that subjects have learned to associate with a reward can disrupt search for a target from a different category, even when the distractors share few visual features with the exemplar rewarded during training (Hickey, Kaiser, & Peelen, 2015). All of these different lines of evidence suggest that CL might involve a change in processing of high-level stimulus information.

A second, independent question about what is learned in CL that was raised by earlier studies is whether CL involves learning efficient processing of the target stimulus, distractors, or both. Some studies suggest that learning to process distractors efficiently (so as to better ignore them) might be more central to the improvements seen in CL than learning to process targets efficiently (Carrasco et al., 1998; Wang et al., 1994). At the same time, there is evidence that memory for distractor stimuli in visual search tends to be poor (Beck, Peterson, Boot, Vomela, & Kramer, 2006). We conducted an additional experiment—Experiment 4—to address this unresolved puzzle.

General methods

The following describes the default experimental parameters of the training task (the “Standard Task”), which were

modified slightly to suit the requirements of each experiment. Any modifications to the Standard Task are described in the [Method](#) section for the individual experiment.

Participants

All participants were Dartmouth College students with normal or corrected-to-normal vision. Participants volunteered to take part in the experiments for payment or course credit under a protocol approved by the local Institutional Review Board. Participants completed seven sessions of training on a visual-search task, followed by one or more test sessions in which a stimulus or task variable was changed. Participants were asked to complete all experimental sessions within 2 weeks. No participant completed more than one of the experiments.

Stimuli, design, and procedure

Standard search stimuli were bisected disks with isoluminant red and green halves. Isoluminance was determined for each participant using flicker photometry, in which the red and green (and, in some experiments, blue and yellow) hues to be presented were shown in rapid alternation (20 Hz) with a neutral color presented at approximately 25% of maximum monitor luminance. Participants adjusted the luminance of the colored stimulus during the presentation to find the point of minimal apparent flicker. Each hue was presented 10 times, with different random starting intensities. Color values for all sessions of the experiment were set at the average point of minimal flicker for eight of the 10 measurements, excluding the first two for each hue as practice trials. Before beginning the color-definition task, all participants were queried for colorblindness, and colorblind prospective participants were not enrolled.

In experiments using the standard red and green disks (the Standard Stimuli), the target was identifiable by a particular configuration of color and position: a red-left, green-right target amongst green-left, red-right distractors. Target and distractor identities remained fixed throughout all experimental sessions for each participant, except when they changed posttraining as part of an experimental manipulation. Stimuli were arranged in concentric rings around a central fixation point, and scaled by eccentricity as a function of cortical magnification factor (R. Duncan & Boynton, 2003). See Fig. 1 for an example display.

The total diameter of the stimulus display subtended 41.6° (degrees of visual angle). The innermost edges of the innermost stimuli were 0.8° from a central fixation point, and those innermost stimuli each subtended 0.6°. Each concentric ring of stimuli was separated from the next by 0.8°. The outermost stimuli each subtended 3.7°.

One hundred target-present trials and 20 target-absent (catch) trials were presented in each experimental session. The number of stimuli presented on each trial varied: 2, 4, 8, 16, or 32 stimuli could appear on any given trial. The frequency of each set size was counterbalanced within each session. White outline rings remained on the screen at all times as location boundaries.

Participants self-initiated each trial. Stimuli remained visible until the participant made a response, shifting slightly within a fraction of the stimulus radius within the white outlines every 100 ms to counteract perceptual fading due to retinal ganglion cell adaptation. Participants' task was to make a five-alternative forced-choice decision about the location of the target as quickly and accurately as possible (i.e., target in ring 1, 2, 3 or 4, or target absent) by pressing one of five corresponding buttons without moving their eyes from the central fixation point. After the end of each trial, participants received feedback about the accuracy of their response via a color change of the fixation point (green = correct, red = incorrect).

Experimental sessions were conducted in a dark room with participants seated 55 cm from a Mitsubishi 2070SB CRT monitor (30 × 40 cm, 1600 × 1200 pixel resolution, 85 Hz refresh rate) driven by a PC running MATLAB (The MathWorks; Natick, MA) Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997).

Behavioral data from the search task were analyzed as follows. Efficiency of visual search within each session was defined by the slope of the search function (i.e., the slope of the line of best fit between the number of search stimuli and median reaction time for that session, for target-present trials), and this measure of search slope was computed for each participant for each session. This analysis yields a standard measure of search efficiency (Wolfe, 1998). Session slopes were adjusted by dividing by the participant's accuracy score for the session (i.e., fast search in a session with low accuracy would produce a slower adjusted score), a standard way of combining reaction time and accuracy measurements (Townsend & Ashby, 1978). These adjusted slopes were used in all subsequent analyses. Because accuracies were high throughout all the experiments, the accuracy-adjustment factors applied were rather small, and no experiment showed any evidence of a speed-accuracy trade-off with learning. Plots showing accuracy as a function of training day and condition for each experiment are provided in the [Supplemental Figures](#).

A single measure of learning rate was computed for each participant by modeling the change over time in adjusted search slope using a power function (i.e., learning rate corresponds to the slope of the log-transformed slopes over time). If the modeled power function had a negative slope (i.e., if search slopes decreased over time), that was considered evidence of learning. The slope of each participant's power function reflected that participant's rate of learning, with more negative slopes corresponding to faster learning.

The effects of posttraining manipulations on performance were assessed using two-tailed repeated-measures statistical tests (*t* tests or ANOVAs, where appropriate) to compare performance in the new condition to the trained condition. Performance in the new condition was compared with performance on the trained condition at the end of training, as well as performance on the first day of training (a measure of initial performance). Effect sizes for repeated-measures *t* tests were calculated as Cohen's *d* using the formula of Rosenthal (1991). ANOVA effect sizes are reported as partial eta squared (η_p^2) for each main effect and interaction.

Preliminary experiment: Assessing task learning and experimental power

Previous studies have shown that CL is not simply task learning by demonstrating that improvements in processing are specific to the learned stimuli (e.g., Reavis et al., 2016). However, this is an important finding to replicate, because if CL effects are attributable largely or entirely to task learning, then learning would be expected to generalize to nearly any stimulus change, undermining the logic of the experiments to come. Thus, the main goal of this preliminary experiment was to confirm that CL is in fact disrupted by a complete change of the stimuli. A secondary goal was to use the effect size of this disruption to estimate the sample size necessary for reasonable power in the novel experiments conducted for this study.

Participants were first trained to search one set of configuration stimuli, then tested with an entirely different type of configuration stimuli, while all task demands were kept constant. Such a complete change of the stimuli has been shown to dramatically impair search performance (Reavis et al., 2016). In this preliminary experiment, we collected data under methodological conditions similar to those of the main experiments in this article as a benchmark measure of the amount of change in performance to be expected when minimal learning transfer occurs. We used this information to calculate the sample size needed to reliably detect a change in performance of similar magnitude for the other experiments described in this article.

Method

Participants A sample of nine students (five females), with a mean age of 19.2 years ($SD = 1.2$), were recruited to participate in the experiment.

Design and procedure Participants trained on the color-shape configurations described in the [General Methods](#) section, then performed a test session of search for a rotated *L* among rotated *T*s, another common type of configuration search with

completely different features (Egeth & Dagenbach, 1991; Kwak, Dagenbach, & Egeth, 1991; Wolfe, 1998).

The color-shape search task (hereafter, R/G, for red/green) was exactly as described in the [General Methods](#). The T/L search task was identical, except for the substitution of *T* distractor stimuli with equal-length line segments conjoined end to center, and an *L*-shaped target with half of the *T* cap missing. The lengths of the long line segments were the same as the diameter of the red and green circles. Each T/L stimulus was presented at a random orientation on each trial; only by chance did a search stimulus occasionally appear at a canonical, letter-like orientation.

All participants trained for 7 days on the R/G-search task. On the eighth day, the stimulus type changed to T/Ls. A significant increase in search slopes between Day 7 and Day 8 would indicate that the CL effect was not entirely attributable to task learning. To estimate an appropriate sample size for subsequent experiments, we performed a power analysis with the effect size from this comparison using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007).

Results and Discussion

Participants' search performance improved significantly between Day 1 and Day 7: The slope of the power functions fitted to each participant's session-wise adjusted search slopes were significantly negative, $t(8) = -3.07$, $p = .02$, $d = 1.02$. Search slopes increased significantly between Day 7 and Day 8, when the search stimuli changed, $t(8) = 3.82$, $p = .005$, $d = 1.27$. Indeed, performance on Day 8 did not differ significantly from the first day of training, $t(8) = -1.64$, $p = .14$, $d = 0.55$. These results are depicted in Fig. 2.

A power analysis of the results showed that to achieve a power of 0.8 in subsequent experiments with similar properties (i.e., an 80% chance of finding a significant change in performance between Day 7 and Day 8 if there is a true performance difference between those 2 days), a sample size of seven participants should be sufficient.

The results of the preliminary experiment demonstrate that CL is not just task learning: When the stimuli are changed, but the task is kept constant, participants' performance suffers significantly. This finding replicates an earlier experiment that tested a similar manipulation (Reavis et al., 2016). This replication confirms that changing learned stimuli can disrupt task performance significantly, which validates the approach of assessing learning transfer that is used in the novel experiments described in this article.

The results of this experiment do not rule out the possibility that some task learning occurs during CL. Although postmanipulation performance was not significantly better than performance on the first day of training, mean search efficiency was numerically better on the day of the manipulation, and it is possible that this difference of means would be

significant in a larger study with greater power. Thus, the results certainly leave open the possibility that *some* task learning may occur in CL, but the significant disruption in performance effected by the change in the search stimuli does effectively rule out the possibility that task learning is the *only* type of learning that occurs, and the large effect size of the disruption suggests that a large fraction of the total learning effect is likely stimulus specific (i.e., perceptual).

The power analysis suggested that for subsequent experiments of similar design, an *N* of 7 would be sufficient to achieve 80% power. Because of concerns about attrition, we aimed to recruit 10 participants per group for the remaining experiments of this study, though it was not always feasible to find a cohort of 10 volunteers due to limitations in the participant pool. At least seven participants completed each of the remaining experiments, meeting the threshold suggested by the power analysis.

Experiment 1a: Retinotopic learning (proximate stimuli)

One possible explanation of the CL effect is that subjects learn to detect or discriminate particular visual features (e.g., color) at particular retinotopic locations. In this way, participants could learn to detect and respond to search targets accurately without needing to process the conjunctive properties of the stimuli at all, by learning that the presence of red (or green) in specific retinotopic locations relative to fixation indicates the presence of a target. Such an explanation of training-dependent improvements in visual search would be in line with the RHT prediction that highly trained observers will utilize task-relevant low-level feature information about the stimuli more effectively, enabling them to perform search more efficiently. This explanation of CL is also in line with many classic findings in the perceptual learning literature, wherein training leads to improvements in the processing of single visual features at trained retinotopic locations (e.g., Fahle et al., 1995; Kami & Sagi, 1991, 1993; Poggio et al., 1992).

To rule out this explanation of CL, we performed an experiment in which participants practiced searching an array of the standard red and green stimuli in which only half of the possible stimulus locations were occupied during training. At test, the stimuli were presented in the unstimulated locations instead. If CL were attributable to improvements in processing stimulus features at specific retinotopic locations, then this manipulation would be expected to cause a significant disruption in search performance.

Method

Participants Ten participants (five female) volunteered to complete the experiment. Participants had an average age of 19.3 years ($SD = 0.9$).

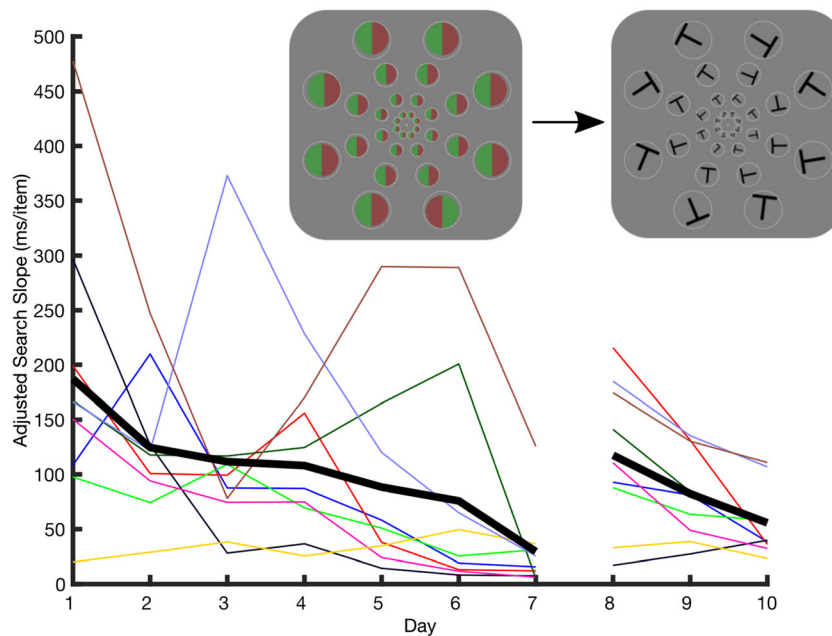


Fig. 2 Effect of stimulus change on search performance. Search slopes declined significantly over 7 days of training with R/G stimuli, but rebounded when search stimuli were changed to T/Ls on Day 8. Each

colored trace represents the data from an individual participant. Heavy black line depicts the group mean. (Color figure online)

Design and procedure The standard visual-search experiment described in the [General Methods](#) section was slightly modified. All of the sizing and spacing parameters were the same as those described in the Standard Task (see the introduction). Four of the eight radii of the visual-search array, however, were never stimulated. Thus, the number of stimuli presented on each trial varied between two and 16. The unstimulated radii alternated spatially with the stimulated radii, and the stimulated and unstimulated radius locations were counterbalanced across participants. All 32 outlines were shown in all sessions. Participants trained for 7 days on one set of radii; on the eighth day, the search stimuli were moved to the other set of radii. A significant increase in search slopes between Day 7 and Day 8 would be indicative of retinotopic learning.

Results

Participants learned to search more efficiently over the 7 days of training: The slope of the power functions describing the change in participants' adjusted search slopes over time was significantly negative, $t(9) = -8.91$, $p < .001$, $d = 2.82$. There was no significant change in participants' adjusted search slopes between Day 7 and Day 8 (i.e., between the old and new retinotopic locations), $t(9) = 0.79$, $p = .44$, $d = 0.25$. Performance on Day 8 was significantly better than on the first day of training, $t(9) = -4.68$, $p = .001$, $d = 1.48$. These results are shown in [Fig. 3](#).

No significant change in performance was evident when the retinotopic position of search stimuli changed between

Day 7 and Day 8, and performance remained significantly better than at the start of training. However, the retinotopic positions of training and test stimuli were in close spatial proximity. Training and test positions in the innermost ring were nearly adjacent, only about 0.1° apart, and 3.8° apart in the outermost ring. Therefore, even small deviations in fixation might have been sufficient to blur the intended separation between training and test locations, causing stimuli to land in “untrained” locations sufficiently often during training that those “untrained” locations were, in fact, trained as well. Experiment [1b](#) was performed to address this issue.

Experiment 1b: Retinotopic learning (distant stimuli)

Experiment [1a](#) showed no significant decline in search performance associated with changing the retinotopic location of stimuli. The training and test locations in Experiment [1a](#), however, were close together. In Experiment [1b](#), the retinotopic specificity of CL was retested using more distant training and test locations where the possibility of overlap would be substantially reduced. All other experimental parameters were kept constant.

Method

Participants Two participants dropped out of the experiment after the first training session and were excluded from all

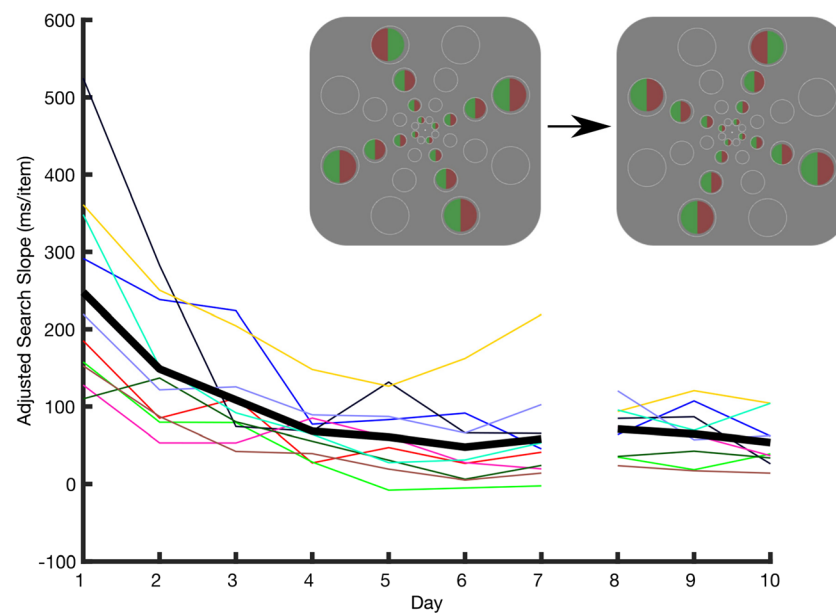


Fig. 3 Experiment 1a: Effect of retinotopic location change on search performance (proximate stimuli). Performance improved significantly over 7 days of training. Changing the retinotopic locations of the search stimuli on Day 8 did not lead to a significant impairment in search

performance. Each colored trace represents the data from an individual participant. Heavy black line represents the group mean. (Color figure online)

analyses. There were nine remaining participants (five females), with a mean age of 20.2 years ($SD = 1.1$).

Design and procedure The methods of Experiment 1b were very similar to Experiment 1a, except that the radii of the search array were moved further into the periphery. Two radii of the usual search array were presented, arranged in an X pattern, farther from fixation. Only two to eight search stimuli were presented per trial. In the new arrangement, the innermost edge of the innermost ring of the stimuli was 5.3° from fixation; the outermost edge of the outermost ring was 17.3° distant. The scaling-by-distance parameter used in the Standard Task (R. Duncan & Boynton, 2003) was applied as before, resulting in slightly larger stimuli, with disks in the innermost ring subtending 1.4° and, in the outermost ring, 4.1° . Because of the more distant arrangement of the search array and the removal of two radii of the array, the training and test radii were 7.0° distant at their most proximate point.

Results

Participants' search efficiency increased significantly over the 7 days of training; Power function slopes fitted to adjusted search slopes over time were significantly negative, $t(8) = -4.78$, $p = .001$, $d = 1.59$. There was a marginally significant increase in search slopes between Day 7 and Day 8, when the retinotopic positions of the stimuli changed: $t(8) = 2.15$, $p = .06$, $d = 0.72$. Performance on Day 8, however, was still significantly better than on Day 1, $t(8) = -2.99$, $p = .02$, $d = 1.00$. These results are plotted in Fig. 4.

Performance suffered marginally when search stimuli moved to new, distant positions on the screen, but remained significantly better after this manipulation than at the beginning of training. Together with the results of Experiment 1a, the results of Experiment 1b suggest that CL is not primarily attributable to improved processing of single visual features at particular retinotopic locations. In both experiments, learned improvements in search efficiency mostly generalized to stimuli moved to untrained locations, as evidenced by participants' significantly better performance on the eighth day of the task than on the first day.

Whereas the marginal postmanipulation performance deficit shown in Experiment 1b leaves open the possibility that some retinotopically specific learning occurs, the relatively high degree of generalization of CL to new retinotopic locations helps to differentiate this type of learning from classical perceptual learning effects, which are often very specific to the retinotopic locations where stimuli are encountered during training. Furthermore, these results fail to support the hypothesis that CL depends mainly on better "vision with scrutiny" whereby participants learn to detect task-relevant single-feature information to perform search efficiently.

Experiment 2: Configural learning (oriented search template)

Another possible explanation of CL is that it involves the development of units tuned selectively to search targets and distractors. This is another prediction of RHT: that expert

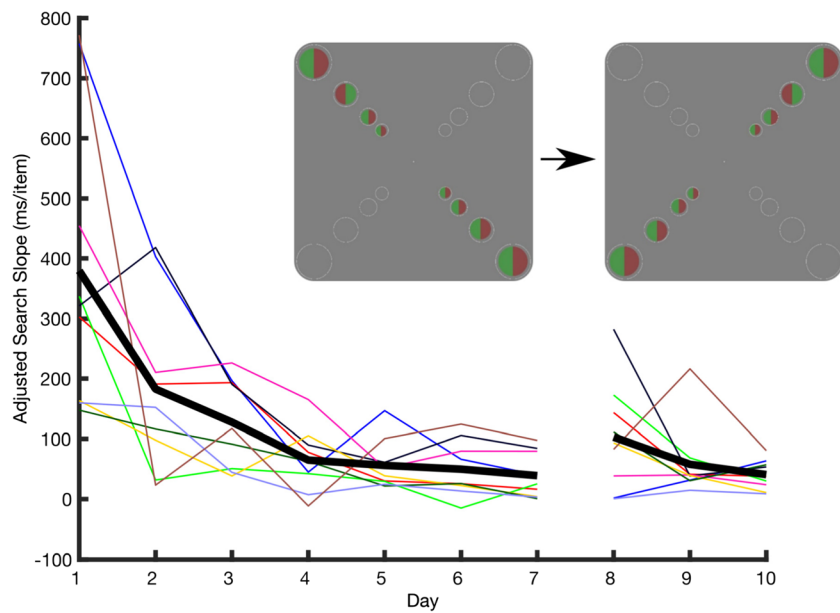


Fig. 4 Experiment 1b: Effect of retinotopic location change on search performance (distant stimuli). Performance on the visual-search task improved significantly over 7 days of training. When the retinotopic locations of the stimuli changed on Day 8, there was a marginally

significant reduction in search performance. Each colored trace represents the data from an individual participant. Heavy black line represents the group mean. (Color figure online)

observers will be able to search arrays of learned stimuli “at a glance,” because units near the top of the visual processing hierarchy will respond selectively to search targets (or distractors). In the case of CL involving the Standard Stimuli, where targets and distractors can be identified by the spatial relationship between two colored elements, such selectivity could be for the relative location of the red versus green elements of the stimulus (e.g., selectivity for red-left, green-right or green-left, red-right stimuli). This possibility is also congruent with theoretical models that propose this type of search would involve matching stimuli to an oriented template representation of the target (Bravo & Farid, 2009, 2012; Chelazzi, Miller, Duncan, & Desimone, 1993; J. Duncan & Humphreys, 1989; Vickery, King, & Jiang, 2005).

If this explanation of CL is correct, then rotating the search stimuli to an intermediate orientation, midway between the learned target and distractor orientations, should cause a significant disruption in performance. This is because newly tuned high-level units should respond equally strongly to the rotated targets and distractors (i.e., both stimuli should match the search template equally well). Experiment 2 tested this explanation of CL.

Participants trained on search arrays containing R/G stimuli in their canonical orientations. After training, participants were tested with arrays of stimuli rotated 90° clockwise and counterclockwise (Day 8 and Day 9, respectively). Experiment 2 included an additional manipulation on the 10th day, which was intended to address a related question. That day, the search stimuli were rotated 180° from the trained orientation (90° from the orientations tested on Day 8 and Day

9). Because of the design of the stimuli, a 180° rotation meant that the learned target and distractor stimuli were presented again on the 10th day, with visual properties identical to those they had during training. The 180° rotation, however, led to an exchange of target and distractor identities such that participants had to search for what had been a target that had been a distractor and ignore distractors that had been targets.

This manipulation was intended to rule out another possible explanation of CL: that it is simply an improvement in the discriminability of targets and distractors. If CL simply involves a change in the perceptual discriminability of the two types of stimuli, then participants’ performance should be unimpaired by the 180° rotation. Previous studies, however, have shown that targets are learned as targets and distractors are learned as distractors, and that exchanging target and distractor identities after this learning has taken place disrupts performance (Shiffrin & Schneider, 1977). Therefore, impairment in performance *was* expected to result from this 180° rotation, and the manipulation was included as an add-on to the experiment to attempt to replicate those earlier findings.

Method

Participants Nine volunteers (six females) completed the experiment. Their mean age was 18.7 years ($SD = 0.9$).

Design and procedure Participants trained on the Standard Task for 7 days. On the eighth day, all stimuli rotated 90° counterclockwise. On the ninth day, all stimuli rotated 90° clockwise from their orientation during training (i.e., 180°

from Day 8 orientations). On the 10th day, all stimuli rotated 180° from their training orientations (i.e., 90° from Day 8 and Day 9).

Results

Participants' search efficiency improved significantly across the 7 days of training, $t(8) = -4.39$, $p = .002$, $d = 1.46$. There was no significant change in search slopes when all stimuli rotated 90° counterclockwise between Day 7 and Day 8, $t(8) = 1.05$, $p = .33$, $d = 0.35$. Likewise, there was no significant difference in performance between Day 8 and Day 9 (90° clockwise rotation from trained orientations, and 180° rotation from Day 8), $t(8) = 0.76$, $p = .47$, $d = 0.25$. There was, however, a large reduction in search efficiency on Day 10 compared with Day 7, when familiar target and distractor identities were reversed (i.e., when stimuli rotated 180°), $t(8) = 5.90$, $p < .001$, $d = 1.97$. Indeed, a post hoc t test comparing search slopes on Day 10 with Day 1 reveals that performance on the last day of the experiment was not significantly better than on the very first day of training, $t(8) = -1.24$, $p = .25$, $d = 0.41$. All effects are shown in Fig. 5.

Stimulus rotations of 90° in either direction from the trained orientation did not significantly disrupt search performance. This result suggests that CL does not depend on the

development of high-level units tuned to the relative spatial position of the red and green stimulus elements (or to the development of simple oriented stimulus templates for target detection).

By contrast, a stimulus rotation of 180°, which is equivalent to an exchange of target and distractor identities, significantly disrupted search performance. In a replication of earlier results, participants' performance returned to pretraining levels of inefficiency following this manipulation. This result is reminiscent of earlier studies wherein practicing visual search with familiar stimuli (e.g., letters) consistently mapped to target and distractor identities led to efficient, automatized visual search that was disrupted by a reversal of the target and distractor identity mappings (Hillstrom & Logan, 1998; Logan, 1988; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Su et al., 2014; Treisman, Vieira, & Hayes, 1992).

The disruption of search performance by a reversal of target and distractor identities in Experiment 2 demonstrates that CL is not attributable solely to an improvement in participants' ability to discriminate target and distractor stimuli defined by the spatial relationship between two colored elements. The significant drop in performance following this manipulation shows that CL also involves learning of the identity of stimuli as targets or distractors.

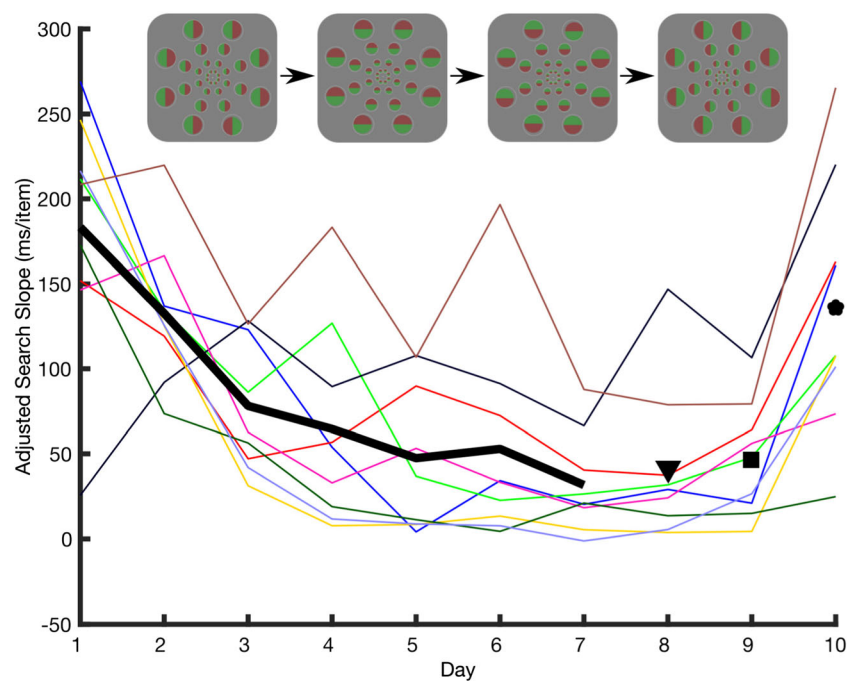


Fig. 5 Experiment 2: Effect of rotation on visual-search performance. Visual-search performance improved significantly over the 7 days of training. After training, participants completed three follow-up test sessions on separate days. On Day 8, all stimuli were rotated 90° counterclockwise from the training orientations; performance was not significantly impaired. On Day 9, stimuli were rotated 90° clockwise from trained orientations (i.e., 180° from Day 8); again, performance

was not significantly impaired. On Day 10, stimuli were rotated 180° from training orientations (i.e., trained target/distractor identities were reversed); here, performance was significantly impaired, and, indeed, not significantly different from performance on the first day of training. Each thin colored trace represents the data from an individual participant. Heavy black line depicts the group mean. (Color figure online)

Experiment 3: Conjunction learning (color-shape template)

Although Experiment 2 ruled out the possibility that CL relies on the development of “detectors” sensitive to the spatial position of the two colored elements, it did not rule out the possibility that subjects could be learning efficient detection of another type of conjunction contained within the search stimuli. In effect, the Standard Stimuli contain three separate conjunctions that are diagnostic of target or distractor identity. Most obviously, they contain the color-position configuration that was the focus of Experiment 2: Targets can be identified as an object with red on the left and green on the right. But the stimuli also contain two diagnostic color-shape conjunctions: The target can be identified by finding the red, left-pointing semicircle, or the green, right-pointing semicircle. Thus, in addition to their configural color-color conjunctions, targets and distractors are identifiable on the basis of two equally informative color-shape conjunctions. If what is actually learned in CL is not the color-color configuration but one or both of the color-shape conjunctions, then perhaps it is not surprising that learning transfers to changes in stimulus orientation. Learning a color-shape conjunction is much like learning an object, and known objects can be identified efficiently at various angles (indeed, the response of object-tuned neurons can be invariant to the orientation of the object; Logothetis, Pauls, & Poggio, 1995).

Experiment 3 was performed to test the hypothesis that participants learn to perceive color-shape conjunctions contained in the standard stimuli more efficiently over time. This was another test of the RHT prediction that CL involves the development of target or distractor “detectors” of some type. This experiment, however, supposed that those detectors might be tuned to different characteristics of the stimuli than those tested in Experiment 2.

The approach employed in Experiment 3 was to perform training using the standard stimuli, then change the color of one of the two constituent color-shape conjunctions, both, or neither. If CL were entirely dependent on the development of a color-shape “detector” (or search template), then learning should completely fail to transfer to a change in the color of both conjunctions (i.e., performance should return to pretraining search efficiency). Because the two color-shape conjunctions contained in each stimulus are equally informative, it is also possible that CL depends on acquisition of a template for just one of them. If this were the case, then a change in one color should completely disrupt performance while a change in the other should not (and a change in both colors should also disrupt performance in this scenario).

Method

Participants Ten participants (six females), with a mean age of 18.2 years ($SD = 0.4$), volunteered to participate.

Design and procedure Before training, all participants were tested using the procedure described in the [General Methods](#) section to determine isoluminant intensities for hues of red, green, blue, and yellow. Then, all participants trained on the standard search task, described in the [General Methods](#) section, for 7 days, with red and green circles. The task was modified such that all trials contained a minimum of four stimuli, so target and distractor identities would always be self-evident.

On the eighth day, all participants were tested in four conditions, which were randomly interleaved in the same session: (1) red/green targets and green/red distractors (no change from training); (2) red/yellow targets and yellow/red distractors (Single-Color Switch I); (3) blue/green targets and green/blue distractors (Single-Color Switch II); (4) blue/yellow targets and yellow/blue distractors (double-color switch).

Results

Participants’ search efficiency improved significantly over the course of the 7 training days, $t(9) = -3.92, p = .004, d = 1.24$. Performance in two conditions on Day 8 was not significantly different from performance on Day 7: the unchanged R/G condition, $t(9) = 1.39, p = .20, d = 0.44$; and the R/Y condition (Single-Switch I), $t(9) = 0.54, p = .60, d = 0.17$. Performance in the other two conditions (Single-Switch II and double switch) was significantly worse on Day 8 than on Day 7: B/G, $t(9) = 3.63, p = .006, d = 1.15$; and B/Y, $t(9) = 5.25, p < .001, d = 1.66$. Performance in all four test conditions on Day 8, however, was still significantly better than performance on Day 1: R/G, $t(9) = -6.91, p < .001, d = 2.19$; B/G, $t(9) = -2.56, p = .03, d = 0.81$; R/Y, $t(9) = -7.94, p < .001, d = 2.51$; B/Y, $t(9) = -4.56, p = .001, d = 1.44$. All effects are shown in Fig. 6.

Experiment 3 shows that CL cannot be entirely explained as a process of developing color-shape conjunction “detectors.” When the colors of both constituent color-shape conjunctions were changed after training (i.e., in the double color switch condition), performance remained significantly better than at the start of training. Nevertheless, two of the color manipulations did cause a significant drop in performance. Both the double-color change (red to blue plus green to yellow) and the single-color change of red to blue (Single-Switch I) significantly disrupted performance of the search task by approximately the same amount.

This result suggests two additional conclusions. First, although CL cannot be completely explained by the acquisition of color-shape conjunction “detectors” or search templates,

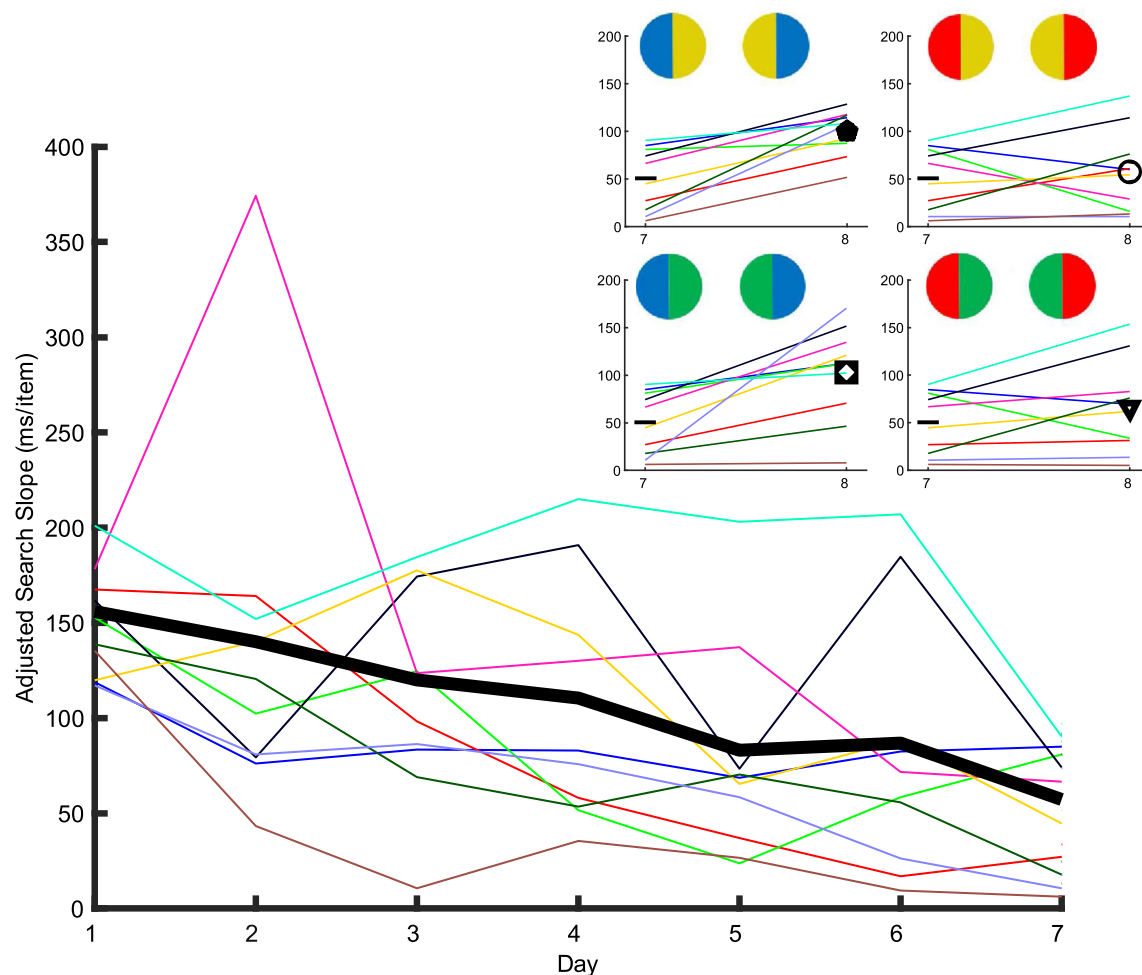


Fig. 6 Experiment 3: Effect of color change on visual-search performance. Visual-search efficiency improved significantly across 7 days of training to find a red/green target among green/red distractors. On Day 8, four types of search stimuli were tested, where one or both colors encountered during training were replaced with a different isoluminant hue. Performance was significantly impaired when the red

hue was replaced with blue, but not when the green hue was replaced with yellow. However, performance in all conditions tested on Day 8 was still better than performance on Day 1. Each thin colored trace represents the data from an individual participant. Heavy black line depicts the group mean. Changes in performance between Day 7 and Day 8 for each condition are plotted in the four breakout panels. (Color figure online)

the results of Experiment 3 imply that learning of color-shape conjunctions is part of CL. Second, it implies that participants might learn just one of the two redundant color-shape cues that define target and distractor identities with equal fidelity. While the statistics suggest that this is often the red semicircle, closer inspection of the data at the individual subject level shows that while most subjects are impaired by the change of red to blue but not green to yellow, some subjects show the opposite pattern (see, e.g., the dark green and teal traces in Fig. 6), while others show impairment only for the double-color change (see the brown trace). The experiment did not include enough subjects to perform subgroup analyses, but in a larger sample it might be possible to identify groups of participants whose performance is more or less impaired by particular color-shape conjunction manipulations.

The sources of individual differences in the precise conjunctions that participants learned, and for the

apparent bias toward learning the red conjunction, are unclear. One very speculative possibility is that the type of conjunction that is learned might be influenced by the focus of participants' attention during training, and that more participants might focus on the red portion of the stimuli than on the green portion (perhaps because red is often used as an attention-grabbing signal in other contexts, like stoplights and editors' pen marks). Other studies that investigated effects of stimulus color on visual search have found similar effects, with participants showing better performance for red targets than targets defined by other colors, suggesting that red stimuli might be particularly salient (Fortier-Gauthier, Dell'Acqua, & Jolicœur, 2013; Lindsey et al., 2010). One such study speculated that the effect might be related to a specialization for the processing of human skin (Lindsey et al., 2010). Nevertheless, the reasons

that observers sometimes show a preference for red stimuli in visual search are not well understood, and this phenomenon remains an intriguing topic for future study.

In summary, Experiment 3 shows that learning of color-shape conjunctions cannot fully account for CL, but that color-shape conjunction learning probably *does* account for some of the CL effect. Furthermore, the particular color-shape conjunction that is learned in search for stimuli that contain several redundant, mutually informative conjunctions may vary from one observer to another, although, in the present data, most participants appeared to learn the red semicircle conjunction.

Experiment 4: Learning of targets versus distractors

Previous studies have suggested that learning to process distractors efficiently (allowing them to be more effectively ignored) might contribute more to CL than learning to process targets efficiently (Carrasco et al., 1998; Wang et al., 1994). Although RHT does not make predictions about the extent to which CL involves learning of target versus distractor stimuli, investigating whether learning is focused on the target, the distractors, or both is a way of addressing the overarching question of what stimulus characteristics participants learn to exploit in CL to perform efficient visual search.

Experiment 4 sought to determine whether a change in the processing of targets, distractors, or both underlies the increase in search efficiency that accompanies CL. In this experiment, we trained participants on a set of targets and distractors that were slightly more complex than the standard stimuli, then introduced a similar, novel stimulus as either a target or a distractor. The logic of the manipulation was that if CL involves a change in target processing, then replacing the target should disrupt search, and replacing the distractors with the novel stimulus should disrupt search if CL involves a change in distractor processing. Those two effects need not be mutually exclusive.

The experiment also included a secondary, orthogonal manipulation with the goal of extending the finding that reversing target and distractor mappings leads to a disruption of search, which was shown on Day 10 of Experiment 2. During the test phase of Experiment 4, the target versus distractor stimulus identities established during training were either maintained or reversed. The manipulation in this experiment offers a purer test of the effect of changing stimulus identities because the other stimulus in each display was always novel, rather than being another trained stimulus as in Experiment 2. Thus, any effects of the change in stimulus identity must relate to a learning effect specific to that stimulus, rather than a conflict between two learned stimulus types. The manipulation of

stimulus identities in Experiment 4 was performed in a 2×2 factorial design with the primary manipulation (replacement of the learned targets or distractors with a novel stimulus).

Method

Participants Nine participants were recruited, but two failed to complete all 8 days of the experiment and were excluded from all analyses. The remaining seven participants (five female) had a mean age of 19.4 years ($SD = 0.8$).

Design and procedure Stimuli were slightly modified from the standard red/green stimuli to allow for additional manipulations. Circles were divided into four quadrants of different colors. All four colors were made isoluminant for each participant using a pretraining flicker-fusion technique (see the [General Methods](#) section). Participants trained for 7 days with fixed target/distractor identities. During training, Y/R/G/B targets (color patterns read clockwise from the upper left) were presented among G/B/Y/R distractors (equivalent to a 180° rotation, as in the Standard Task). On the eighth day, a novel stimulus was introduced: B/Y/R/G (equivalent to a 90° rotation from either the target or distractor orientation). This novel stimulus always replaced one of the two learned stimuli, and participants were tested in four conditions, randomly interleaved, following a 2×2 factorial design. Either learned targets or learned distractors were replaced with the novel stimulus. On half of the trials, the trained stimulus retained its trained identity, and on the other half, its target/distractor identity was reversed. Differences in performance between the four conditions tested on Day 8 were identified using a 2×2 repeated-measures ANOVA. Differences between performance in each of those conditions and performance on the last day of training were investigated using four paired-samples t tests.

Results

The results of Experiment 4 are shown in Fig. 7. Participants' performance on the search task improved significantly over the 7 training days, $t(6) = -3.26$, $p = .02$, $d = 1.23$. On the eighth day, performance worsened significantly in all four conditions, relative to performance on Day 7: trained target among novel distractors, $t(6) = 12.10$, $p < .001$, $d = 4.57$; novel target among trained distractors, $t(6) = 3.99$, $p = .007$, $d = 1.51$; trained distractor as a target among novel distractors, $t(6) = 8.04$, $p < .001$, $d = 3.04$; novel target among trained targets as distractors, $t(6) = 7.48$, $p < .001$, $d = 2.83$. On Day 8, performance with the trained target among novel distractors was significantly better than performance on the first day of training, $t(6) = -2.46$, $p = .049$, $d = 0.93$, but performance in the other three conditions did not differ significantly from performance on Day 1 (novel target among trained distractors,

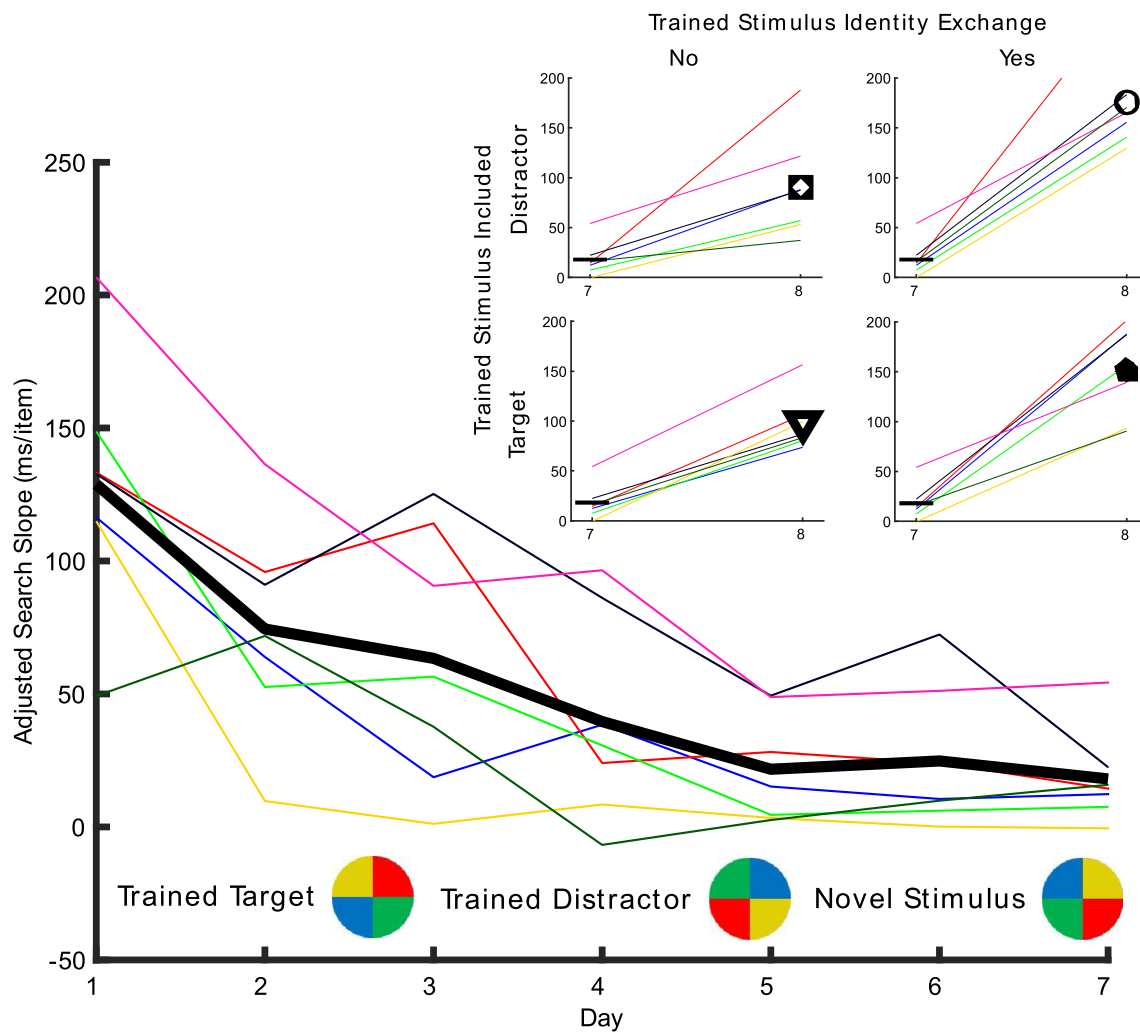


Fig. 7 Experiment 4: Effect of target/distractor replacement with a novel stimulus. Performance improved significantly over 7 days of visual-search training. On Day 8, four different experimental conditions were introduced, wherein one type of stimulus (learned target or distractor) was replaced with a novel one, and learned target/distractor identities were either preserved or reversed. Performance was significantly impaired in

all four conditions, relative to the end of training, but it was most impaired in the two conditions where target and distractor identities were reversed. Each thin colored trace represents the data from an individual participant. Heavy black line depicts the group mean. Changes in performance between Day 7 and Day 8 for each condition are shown in the four breakout panels. (Color figure online)

$t(6) = -2.04, p = .09, d = 0.77$; trained distractor among novel distractors, $t(6) = 1.80, p = .12, d = 0.68$; novel target among trained targets, $t(6) = 1.13, p = .30, d = 0.43$.

Performance in the four conditions tested on the eighth day was compared with a 2×2 repeated-measures ANOVA. Factors were change of trained mapping (yes/no), and trained stimulus type tested (target/distractor). There was a significant main effect of changes to trained mappings, $F(1, 6) = 31.27, p = .001, \eta_p^2 = 0.84$, but no main effect of trained stimulus type tested, $F(1, 6) = 0.43, p = .54, \eta_p^2 = 0.07$, and no significant interaction, $F(1, 6) = 2.05, p = .20, \eta_p^2 = 0.25$.

The results of Experiment 4 support two major conclusions. First, this type of CL involves changes in processing of both target and distractor stimuli. This conclusion comes from the finding that replacing either targets or distractors with a novel stimulus leads to a significant drop in search

efficiency. Second, Experiment 4 shows that changing the target versus distractor mapping of either type of stimulus disrupts learned search efficiency. This result replicates and extends the effect found on day ten of Experiment 2, showing that the disruption caused by exchanging target and distractor identities does not require the creation of a conflict between both types of learned stimuli. Rather, changing the mapping of either targets or distractors is sufficient to disrupt performance.

Together, these findings suggest that targets are learned *as targets* and distractors are learned *as distractors*; the learning effect is not merely attributable to an increase in the discriminability of the two types of stimuli. Our results are consistent with previous studies that have shown with different types of stimuli that practicing visual search with stimuli consistently categorized as targets and distractors can lead to improved task “automaticity” (Czerwinski, Lightfoot, & Shiffrin,

1992; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Treisman et al., 1992). We suspect that the automaticity we observe is related to the learning of stimulus–response associations. Thus, when participants must suddenly change their response because a target is no longer a target, performance suffers. This explanation could be tested more explicitly in a future study by manipulating the response required to stimuli after CL, rather than changing the stimuli themselves.

General conclusions

In a series of experiments, we investigated what stimulus characteristics participants learn to use to perform efficient visual search of complex stimuli containing diagnostic spatial configurations and color-shape conjunctions. Together, the experiments show that performance improvements involve a form of learning of both target and distractor stimuli that cannot be explained *solely* as a process of learning to identify diagnostic features at particular retinotopic locations, acquiring an inflexible selective representation of color-shape conjunctions or the relational position of stimulus elements that differentiates targets and distractors (akin to learning a search template), or merely improving the discrimination of targets from distractors. None of these simple explanations suffices to explain this type of CL because participants continued to process search stimuli more efficiently than they did at the start of training when we manipulated stimulus properties linked to these hypotheses. The overall pattern of learning transfer and transfer failures in the experiments appear consistent with two potential explanatory frameworks for this type of CL.

One remote possibility, in keeping with RHT, is that subjects acquire some type of ineffable high-level representations of targets and distractors that are selective for the different types of stimuli yet broadly tuned enough to remain selective for targets and distractors (and not to switch their responsivity to targets versus distractors) even when major transformations of the stimuli are applied (e.g., 90° rotation in either direction or color changes). While theoretically possible, this seems an unlikely explanation and it is difficult to imagine what tuning properties would be required for such units to lead to the pattern of transfer and transfer failures we found.

A second possibility, which is much more compelling, is that this type of CL is neither a high-level nor a low-level learning process, as proposed by traditional theories of perceptual learning like RHT. Rather, CL for stimuli like these might involve an emergent phenomenon that reflects many different learning processes at different levels of visual processing. While it is clear that this form of CL cannot be explained as *just* an effect of learning color-shape conjunctions or oriented templates or retinotopic feature detection, nor can it be explained as merely task learning, our data are consistent with the hypothesis that CL involves *all of those things* and

more, at least in some cases. Although changing stimulus parameters related to any one of those factors did not abolish learned improvements in processing, various changes each produced a small to moderate reduction in performance. According to this explanation, although learning to utilize some types of information (e.g., about color-shape conjunctions) might account for a greater fraction of the overall learning effect than learning to utilize other types of information (e.g., retinotopic position of individual features), subjects would be expected to learn to utilize information from various available sources to perform efficient search. Thus, according to this framework, CL might be thought of as an emergent phenomenon that reflects the sum of many small learning effects.

This complex, multilevel account of CL is in line with current models of visual search and perceptual learning. It is consistent with the guided-search model, which proposes that visual search for a target can be guided by various types of information that are combined to form a salience map, which can guide attention to possible target locations in order of descending salience (Wolfe, Cave, & Franzel, 1989; Wolfe & Horowitz, 2017). It is also consistent with recent theoretical and computational models of perceptual learning, which propose that representation and read-out of task-relevant information encoded at many levels of the visual hierarchy improve with training (Doshier & Lu, 2017; Maniglia & Seitz, 2018). Furthermore, if the improvements in search efficiency that accompany CL involve strengthening associations between task-relevant information and consistently mapped responses (i.e., stimulus–response mappings), then this could help to explain the impairments in performance that arise when target or distractor identities are reversed after training.

A formal test of this multilevel explanation of CL using a computational modeling approach, which would permit estimation of the relative influence of different learning effects, would unfortunately require data from more participants than we collected in the present study. While the results of our experiments appear consistent with this account of CL, a quantitative test of the explanatory power of this framework will be an important issue for future research. Thus, the small experimental samples we tested are a major limitation of this study.

Another important limitation is use of one type of stimulus (bisected disks) for most of the experiments. These stimuli are somewhat atypical in the sense that they contain both color-shape and color-position conjunctions that are diagnostic of target and distractor identities. It is possible that CL follows different rules for other types of conjunction stimuli that contain fewer diagnostic conjunctions, or different types of feature conjunctions. Furthermore, because of the design of the stimuli in the present study, it was not possible to determine conclusively to what extent the observed CL effects depended on learning of the color-shape versus color-position

conjunctions. One way in which this issue could be explored in the future would be to repeat experiments like the ones in this study using bisected square stimuli (similar to Heathcote & Mewhort, 1993, and other seminal studies). Because the shape of the two halves of the stimuli would be identical, no diagnostic color-shape conjunctions would be present, and any CL effects would be solely attributable to learning of color-position (i.e., configural) conjunctions.

Although there is clearly still work to be done to fully characterize CL and the mechanisms by which it occurs for different types of stimuli, the present results narrow down the possible explanations of CL for one class of conjunction stimuli. Our findings challenge the notion that this type of CL can be reduced to one simple learning effect that happens to occur with complex stimuli. Rather, the present results suggest it might be an emergent phenomenon that reflects multiple constituent learning effects, which capitalize on different types of information contained in the stimuli. This helps to explain the finding that CL of this type appears to be far more generalizable than many classical perceptual learning effects; many changes to the stimuli might affect the availability of some types of diagnostic information, but not other types. In sum, CL remains an intriguing type of learning that could be an important means of perceptual optimization in ecological vision, and it remains an important topic for further research.

Acknowledgements The authors wish to thank numerous undergraduate research assistants for their help with data collection. E.R. was supported by a National Science Foundation graduate research fellowship during data collection for this project (DGE-1313911). The research was supported by internal Dartmouth funding, Templeton Foundation Grant 14316 to P.T. and National Science Foundation Grant 1632738 to P.T. E.R. is currently supported by a postdoctoral Ruth L. Kirschstein National Research Service Award from the National Institutes of Health (F32MH108317).

References

- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. doi:<https://doi.org/10.1016/j.tics.2004.08.011>
- Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*, 285–299. doi:<https://doi.org/10.1098/rstb.2008.0253>
- Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, *218*(4573), 697–698.
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953–965.
- Beck, M. R., Peterson, M. S., Boot, W. R., Vomela, M., & Kramer, A. F. (2006). Explicit memory for rejected distractors during visual search. *Visual Cognition*, *14*(2), 150–174. doi:<https://doi.org/10.1080/13506280600574487>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.
- Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision*, *9*(1), 1–9. doi:<https://doi.org/10.1167/9.1.34>
- Bravo, M. J., & Farid, H. (2012). Task demands determine the specificity of the search template. *Perception & Psychophysics*, *74*(1), 124–131. doi:<https://doi.org/10.3758/s13414-011-0224-5>
- Carrasco, M., Ponte, D., Rechea, C., & Sampedro, M. J. (1998). “Transient structures”: The effects of practice and distractor grouping on within-dimension conjunction searches. *Perception & Psychophysics*, *60*(7), 1243–1258.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*, 345–347.
- Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *The American Journal of Psychology*, *105*(2), 271–315.
- Doshier, B., & Lu, Z.-L. (2017). Visual perceptual learning and models. *Annual Review of Vision Science*, *3*, 343–363.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458.
- Duncan, R., & Boynton, G. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*, 659–671.
- Egeth, H., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 551–560.
- Ellison, A., & Walsh, V. (1998). Perceptual learning in visual search: Some evidence of specificities. *Vision Research*, *38*(3), 333–345.
- Fahle, M. (1997). Specificity of learning curvature, orientation, and vernier discriminations. *Vision Research*, *37*(14), 1885–1895.
- Fahle, M. (2004). Perceptual learning: A case for early selection. *Journal of Vision*, *4*, 879–890. doi:<https://doi.org/10.1167/4.10.4>
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, *35*(21), 3003–3013.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, *6*(3), 292–297.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.
- Fortier-Gauthier, U., Dell’Acqua, R., & Jolicoeur, P. (2013). The “red-alert” effect in visual search: Evidence from human electrophysiology. *Psychophysiology*, *50*, 671–679. doi:<https://doi.org/10.1111/psyp.12050>
- Frank, S. M., Greenlee, M. W., & Tse, P. U. (2018). Long time no see: Enduring behavioral and neuronal changes in perceptual learning of motion trajectories 3 years after training. *Cerebral Cortex*, *28*(4), 1260–1271. doi:<https://doi.org/10.1093/cercor/bhx039>
- Frank, S. M., Reavis, E. A., Greenlee, M. W., & Tse, P. U. (2016). Pretraining cortical thickness predicts subsequent perceptual learning rate in a visual search task. *Cerebral Cortex*, *26*, 1–10. doi:<https://doi.org/10.1093/cercor/bhu309>
- Frank, S. M., Reavis, E. A., Tse, P. U., & Greenlee, M. W. (2014). Neural mechanisms of feature conjunction learning: Enduring changes in occipital cortex after a week of training. *Human Brain Mapping*, *35*(4), 1201–1211. doi:<https://doi.org/10.1002/hbm.22245>
- Gibson, E. (1963). Perceptual learning. *Annual Review of Psychology*, *14*, 29–56.
- Gold, J. I., & Watanabe, T. (2010). Perceptual learning. *Current Biology*, *20*(2), 46–48.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612. doi:<https://doi.org/10.1146/annurev.psych.49.1.585>

- Harris, H., Glicksberg, M., & Sagi, D. (2012). Generalized perceptual learning in the absence of sensory adaptation. *Current Biology*, 22(19), 1813–1817. doi:<https://doi.org/10.1016/j.cub.2012.07.059>
- Heathcote, A., & Mewhort, D. J. K. (1993). Representation and selection of relative position. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 488–516.
- Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General*, 144(2), 264–273.
- Hillstrom, A. P., & Logan, G. D. (1998). Decomposing visual search: Evidence of multiple item-specific skills. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5), 1385–1398.
- Hochstein, S., & Ahissar, M. (2002). View from the top : Hierarchies and reverse hierarchies review. *Neuron*, 36, 791–804.
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, 88(11), 4966–4970.
- Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, 365(6443), 250–252. doi:<https://doi.org/10.1038/365250a0>
- Kwak, H. W., Dagenbach, D., & Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception & Psychophysics*, 49(5), 473–480.
- Lindsey, D. T., Brown, A. M., Reijnen, E., Rich, A. N., Kuzmova, Y. I., & Wolfe, J. M. (2010). Color channels, not color appearance or color categories, guide visual search for desaturated color targets. *Psychological Science*, 21(9), 1208–1214. doi:<https://doi.org/10.1177/0956797610379861>
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492–527.
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology: CB*, 5(5), 552–563.
- Maniglia, M., & Seitz, A. R. (2018). Towards a whole brain model of perceptual learning. *Current Opinion in Behavioral Sciences*, 20, 47–55. doi:<https://doi.org/10.1016/j.cobeha.2017.10.004>
- Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, 256(5059), 1018–1021.
- Reavis, E. A., Frank, S. M., Greenlee, M. W., & Tse, P. U. (2016). Neural correlates of context-dependent feature-conjunction learning in visual search tasks. *Human Brain Mapping*, 37, 2319–2330.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: SAGE.
- Sasaki, Y., Nájuez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, 11(1), 53–60. doi:<https://doi.org/10.1038/nrn2737>
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84(1), 1–66.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84(2), 127–190.
- Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics*, 52(5), 582–588.
- Sireteanu, R., & Rettenbach, R. (1995). Perceptual learning in visual search: Fast, enduring, but non-specific. *Vision Research*, 35(14), 2037–2043.
- Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Research*, 40(21), 2925–2949.
- Sripati, A. P., & Olson, C. R. (2010). Global image dissimilarity in macaque inferotemporal cortex predicts human visual search efficiency. *Journal of Neuroscience*, 30(4), 1258–1269. doi:<https://doi.org/10.1523/JNEUROSCI.1908-09.2010>
- Su, Y., Lai, Y., Huang, W., Tan, W., Qu, Z., & Ding, Y. (2014). Short-term perceptual learning in visual conjunction search. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1415–1424. doi:<https://doi.org/10.1037/a0036337>
- Townsend, J., & Ashby, F. (1978). Methods of modeling capacity in simple processing systems. In N. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 199–239). Hillsdale, NJ: Erlbaum.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *The American Journal of Psychology*, 105(2), 341–362.
- Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, 5, 81–92. doi:<https://doi.org/10.1167/5.1.8>
- Walsh, V., Ashbridge, E., & Cowey, A. (1998). Cortical plasticity in perceptual learning demonstrated by transcranial magnetic stimulation. *Neuropsychologia*, 36(4), 363–367.
- Wang, Q., Cavanagh, P., & Green, M. F. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, 56(5), 495–500.
- Wang, R., Wang, J., Zhang, J.-Y., Xie, X.-Y., Yang, Y.-X., Luo, S.-H., ... Li, W. (2016). Perceptual learning at a conceptual level. *Journal of Neuroscience*, 36(7), 2238–2246. <https://doi.org/10.1523/JNEUROSCI.2732-15.2016>
- Wang, R., Zhang, J., Klein, S. A., & Levi, D. M. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A “piggybacking” effect. *Journal of Vision*, 14, 1–10. doi:<https://doi.org/10.1167/14.13.12>
- Watanabe, T., Nájuez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, 413(6858), 844–848. doi:<https://doi.org/10.1038/35101601>
- Watanabe, T., & Sasaki, Y. (2015). Perceptual learning : Toward a comprehensive theory. *Annual Review of Psychology*, 66(August), 197–221. doi:<https://doi.org/10.1146/annurev-psych-010814-015214>
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention*. London, UK: University College London Press.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 419–433.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behavior*, 1(58), 1–8. doi:<https://doi.org/10.1038/s41562-017-0058>
- Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance*, 39(3), 861–871. doi:<https://doi.org/10.1037/a0030517>
- Xiao, L.-Q., Zhang, J.-Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, 18(24), 1922–1926. doi:<https://doi.org/10.1016/j.cub.2008.10.030>
- Yashar, A., & Carrasco, M. (2016). Rapid and long-lasting learning of feature binding. *Cognition*, 154, 130–138. doi:<https://doi.org/10.1016/j.cognition.2016.05.019>